Warwick-Edinburgh Mental Well-being Scale
(WEMWBS)

User Guide
Version 1

Professor Sarah Stewart-Brown
Professor of Public Health
sarah.stewart-brown@warwick.ac.uk
&
Dr Kulsum Janmohamed
Academic Clinical Fellow
k.janmohamed@warwick.ac.uk

Warwick Medical School
University of Warwick

Edited by Dr Jane Parkinson
Public Health Adviser
NHS Health Scotland

June 2008

Acknowledgements

The Warwick-Edinburgh Mental Well-being Scale was funded by the Scottish Government's[1] National Programme for Improving Mental Health and Well-being, commissioned by NHS Health Scotland, developed by the University of Warwick and the University of Edinburgh, and is jointly owned by NHS Health Scotland, the University of Warwick and the University of Edinburgh.

Acknowledgements go to the following:

- Members of the research team involved in the development and validation of WEMWBS: Ruth Fishwick (University of Warwick), Louise Hiller (University of Warwick), Professor Stephen Joseph (University of Nottingham), Professor Stephen Platt (University of Edinburgh), Professor Sarah Stewart-Brown (University of Warwick) and Ruth Tennant (University of Warwick).
- Members of the research Advisory Group: Professor Glynn Lewis (University of Bristol), Dr Jane Parkinson (NHS Health Scotland), Professor Jenny Secker (Anglia Ruskin University), Professor Stephen Stansfeld (University of London) and Professor Scott Weich (University of Warwick).
- Those who commented on this manual: Dr David Gordon (NHS Health Scotland), Emma Hogg (NHS Health Scotland), Kate O'Hara (CSIP West Midlands), and Professor Stephen Platt (University of Edinburgh).

Audience

It is anticipated that the audience for this manual includes researchers and practitioners who are familiar with the use of scales in evaluations. This manual does not seek to answer questions relating to what to consider for evaluation purposes. For this, the evaluation guides of NHS Health Scotland provide the required information and should be referred to in the first instance (www.healthscotland.com/mental-health-publications.aspx).

Update revisions

This manual will be updated and revised as necessary as further validation and data on WEMWBS become available. For the current version at any time see http://www.healthscotland.com/scotlands-health/population/Measuring-positive-mental-health.aspx.

[1] Previously known as the Scottish Executive

Summary

The Warwick-Edinburgh Mental Well-being Scale (WEMWBS) was developed by researchers at the Universities of Warwick and Edinburgh, with funding provided by NHS Health Scotland, to enable the measurement of mental well-being of adults in the UK.

WEMWBS is a 14 item scale of mental well-being covering subjective well-being and psychological functioning, in which all items are worded positively and address aspects of positive mental health. The scale is scored by summing responses to each item answered on a 1 to 5 Likert scale. The minimum scale score is 14 and the maximum is 70. WEMWBS has been validated for use in the UK with those aged 16 and above. Validation involved both student and general population samples, and focus groups.

People participating in studies of face validity found the scale clear, unambiguous and easy to complete. They volunteered the opinion that the scale measured mental well-being.

Population scores on WEMWBS approximate to a normal distribution with no ceiling or floor effects, making the scale suitable for monitoring mental well-being in population samples. The scale is not designed to identify individuals with exceptionally high or low positive mental health, so no 'cut off' has been developed (analogous to a mental illness 'cut-off' on for example the GHQ 12 scale). The provisional Scottish population mean score is 50.7 with a 95% confidence interval of 50.3 to 51.1, obtained from a combined national dataset comprising data from the *Health Education Population Survey* 2006 (wave 12) and the *Well? What do you think?* 2006 survey.

Scores derived from the student and population samples show a single underlying factor, interpreted to be mental well-being, with low levels of social desirability bias and expected moderate correlations with other scales of well-being. Scores for individuals are stable over a one week period.

In general population samples, significant differences in WEMWBS scores were found by certain factors such as tenure, employment status, and marital status. Non-significant trends were found between mental well-being and social grade (with lowest scores among those in the most deprived groups), a u-shaped relationship was found for age and small but non-significant differences were found for sex (male scores were slightly higher).

Further research on WEMWBS is ongoing. This includes: establishing WEMWBS's sensitivity to change; assessing its 'scaling properties' and the potential to reduce the number of items; and validation to determine whether WEMWBS can be used with children aged 13 to 15 years of age. Other research still required includes assessing the extent to which it is appropriate to use WEMWBS to assess mental well-being in English speaking ethnic minority populations in the UK.

As a short and psychometrically robust scale, with no ceiling effects in population samples, WEMWBS offers promise as a tool for monitoring mental well-being at a population level. It is freely available but prospective users should register with Dr Kulsum Janmohamed K.janmohamed@warwick.ac.uk or Professor Sarah Stewart-Brown sarah.stewart-brown@warwick.ac.uk. If the scale is reproduced it must remain unaltered and include the copyright statement which appears with it (Appendix ii).

Contents

1. Introduction 1

2. A word about mental well-being 2

3. What is WEMWBS and how was it developed? 3

4. Validation of WEMWBS 4

Student populations 4

Focus groups 7

Scottish population samples 8

5. Distribution of WEMWBS scores 9

6. Comparison between WEMWBS scores and scores on the GHQ 12 11

7. Variation across demographic & social groups 13

8. Using WEMWBS 15

9. Current usage of WEMWBS 18

10. Further validation research on WEMWBS 19

Bibliography 21

Appendix i Conditions of using WEMWBS 22

Appendix ii Description of scales used to assess the construct validity of WEMWBS 24

Appendix iii WEMWBS median scores across demographic groups 26

1. Introduction

Practitioners of mental health promotion and public mental health have for many years recognised the need to focus their efforts on improving mental health as well as preventing mental illness. Because of confusion relating to use of the term 'mental health' to describe services for people with mental illness, terms like positive mental health and mental well-being have been adopted to describe these initiatives Positive mental health and mental well-being are used interchangeably in this manual).

Efforts to promote mental well-being have been hampered by a lack of valid instruments which are suitable for measuring these attributes in the general population. The monitoring of population mental well-being and the evaluation of interventions to promote positive mental health has therefore had to be undertaken using instruments designed primarily to detect mental illness. There are two problems with such an approach. First, mental illness measures tend to have significant ceiling effects in general population samples, meaning that people with only moderately good mental health can achieve the highest possible score. As a result the instrument cannot show improvements in mental health in the healthier portion of the population distribution. Second, participants who are involved in the evaluation of interventions to promote mental health may develop the erroneous impression that the interventions are designed only to help people with mental health problems and in this way the evaluation can affect the impact of interventions.

To overcome these problems NHS Health Scotland commissioned the development of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) as part of the Mental Health Indicators Programme.[2]

This manual is for those who want to use WEMWBS for monitoring and research purposes as well as for evaluations. Those who require information on what to consider for evaluation are referred to the NHS Health Scotland evaluation guides in the first instance (www.healthscotland.com/mental-health-publications.aspx).

Ceiling and floor effects – these occur when many people score the maximum or minimum score on a scale. Improvements or deteriorations in the assessed variable being measured cannot therefore be identified. For example, significant ceiling effects in a mental health scale used in a general population sample may mean that people who possess only moderately good mental health can achieve the highest possible score. As a result the instrument cannot show improvements in mental health in the healthier portion of the population distribution.

[2] NHS Health Scotland was commissioned by the Scottish Government's National Programme for Improving Mental Health and Well-being (www.wellscotland.info) to establish a core set of national, sustainable mental health and well-being indicators for adults in Scotland (www.healthscotland.com/understanding/population/mental-health-indicators.aspx).

2. A word about mental well-being

A necessary starting point for the development of a new instrument is a clear understanding of the concept which it is designed to measure. In the past there has been considerable discussion and debate about the nature of positive mental health and well-being. Recently a reasonable level of consensus has emerged among both academics and among the public.

Mental well-being is now largely accepted as covering two perspectives: (1) the subjective experience of happiness (affect) and life satisfaction (the hedonic perspective); and (2) positive psychological functioning, good relationships with others and self realisation (the eudaimonic perspective). The latter includes the capacity for self development, positive relations with others, autonomy, self acceptance and competence. Those wanting to understand more about this subject are referred to the large literature, clearly described in Ryan and Deci (2001).

There has been some discussion in the academic literature as to whether mental well-being and mental illness represent two ends of a single spectrum (single continuum model) or two separate dimensions (two continua or dual continua model). The two continua model allows for the possibility that people who have mental illnesses can experience mental well-being. It reflects the finding that analysis of instruments covering both positive and negative mental health often suggests two correlated but independent underlying factors. Possible explanations for these findings include issues relating to how psychiatric conditions are defined, the fluctuating nature of mental illness, and individuals' interpretations and responses to positively and negatively worded items on mental health measurement scales.

Mental well-being relates to a person's psychological functioning, life-satisfaction and ability to develop and maintain mutually benefiting relationships. Psychological well-being includes the ability to maintain a sense of autonomy, self acceptance, personal growth, purpose in life and self esteem. Staying mentally healthy is more than treating or preventing mental illness.

Mental illness is a term to encompass mental disorders – these are illnesses which affect mood, affect and the ability to function effectively and appropriately.

Hedonic perspective of well-being focuses on the subjective experience of happiness (affect) and life satisfaction.

Eudaimonic perspective of well-being focuses on psychological functioning, good relationships with others and self realisation. This is the development of human potential which when realised results in positive functioning in life, and covers a wide range of cognitive aspects of mental health.

3. What is WEMWBS and how was it developed?

The Warwick-Edinburgh Mental Well-being Scale (WEMWBS) comprises 14 items that relate to an individual's state of mental well-being (thoughts and feelings) in the previous two weeks (see Appendix i). Responses are made on a 5-point scale ranging from 'none of the time' to 'all of the time'. Each item is worded positively and together they cover most, but not all, attributes of mental well-being including both hedonic and eudaimonic perspectives. Areas not covered include spirituality or purpose in life. These were deemed to extend beyond the general population's current understanding of mental well-being and their inclusion was thought likely to increase non-response.

WEMWBS aims to measure mental well-being itself and not the determinants of mental well-being, which include resilience, skills in relationship, conflict management and problem solving, as well as socioeconomic factors such as poverty, domestic violence, bullying, unemployment, stigma, racism and other forms of social exclusion.

WEMWBS was developed through research that was conducted at Warwick and Edinburgh Universities. The starting point for the research was a pre-existing scale called the Affectometer 2, developed in the 1980s in New Zealand (Kammann & Flett, 1983). Affectometer 2 consists of 20 statements and 20 adjectives relating to mental health in which positive and negative items are balanced. It proved to have a broad measure of intuitive appeal to practitioners and researchers working in this area in the UK. While it had been used in a number of countries, there was no UK validation of the scale and so this validation was conducted as the first step in this research project.

Validation of the Affectometer 2 in both population and student samples suggested that whilst it performed adequately, it was longer than need be and subject to an unacceptable level of bias due to 'desirable responding' (respondents answering in a way they thought was likely to be 'approved of') (Tennant *et al.,* 2006; Tennant, Joseph & Stewart-Brown, 2007). A focus group study involving participants from a wide range of socio-economic backgrounds found that although in general the scale was viewed favourably, some of the items were considered to be 'difficult', and in spite of the balance of positive and negative items, the instrument was viewed predominantly as a measure of mental illness (Tennant *et al*., 2006).

These results were reviewed by a multidisciplinary research advisory group familiar with epidemiological research as well as the academic literature relating to concepts of positive mental health. The research team drafted a set of items derived partly from Affectometer 2, but taking into account the findings of the qualitative focus group research relating to difficult and potentially redundant items, whilst at all times referring to current literature on positive mental health. Working iteratively with members of the advisory group this new scale was refined to the 14 item scale WEMWBS.

4. Validation of WEMWBS

Validation to date has been performed in the UK with those aged 16 and above. WEMWBS was initially validated in student samples recruited at the universities of Warwick and Edinburgh in 2006, and subsequently discussed by two mini-focus groups in Scotland and England (Tennant *et al*., 2006; Tennant *et al*., 2007). WEMWBS was then included in two national Scottish population surveys in 2006 allowing validation using population data.

Table 1 below lists whether or not the psychometric tests involved in validating a scale have been performed on WEMWBS and if so the sample(s) used. Details of the results are given on the following pages.

Table 1: Psychometric testing of WEMWBS

Psychometric test	Tested	Sample
Principal components factor analysis	✓	Student population samples & Scottish general population samples
Construct validity	✓	Student population samples & Scottish general population samples
Internal consistency	✓	Student population samples & Scottish general population samples
Test-retest reliability	✓	Student population samples
Response Bias	✓	Student population samples
Face (or content) validity	✓	WEMWBS research advisory group & Focus groups
Rasch analysis	✓	Scottish general population samples
Sensitivity to change	x	Currently being assessed
Criterion validity	x	'Gold standard' measure to assess WEMWBS against does not currently exist
Cross-cultural validity	x	Interest has been expressed in using WEMWBS in other countries. An Icelandic version has been created by translation and back-translation.

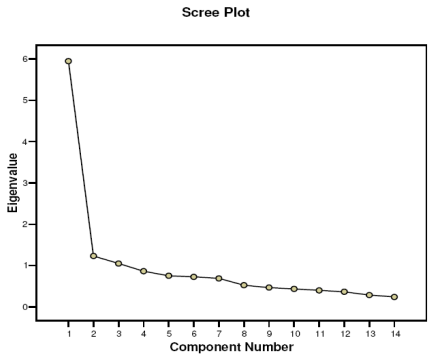
Student populations (n = 348)

Principal components factor analysis
The main aims in conducting these analyses were:

- to determine whether the number of variables in the scale can be reduced
- to determine the relationships between variables

The test considers how much variance is added by each factor the scale considers. The variance that each additional factor contributes is expressed by eigenvalues. A scree plot of eigenvalues against component numbers can be used to illustrate the amount of variance that a single factor contributes.

Figure 1: WEMWBS Scree Plot for student samples (n = 348)



Factor analysis confirmed a single underlying factor to the scale, shown in the sharp 'elbow' of the screen plot. This underlying factor is interpreted to be mental well-being.

Construct validity

Considers the extent to which there are logical relationships between the scale and other scales or factors known to affect the concept being measured (such as age or sex). It is assessed by correlations between the scale under review and other scales measuring similar concepts (convergent validity) or different concepts (divergent validity) and by determining statistically significant differences in scale scores between different groups.

For the validation of WEMWBS, this was assessed by testing correlations between WEMWBS and other scales that measure aspects of mental health, as well as scales that measure general health and emotional intelligence (Appendix ii), and also the extent to which it follows anticipated patterns for age and sex.

Table 2: Correlation of WEMWBS to other scales

Scale	n	Correlation with WEMWBS$^{\alpha}$
WHO-Five Well-being Index	79	0.77**
Short Depression Happiness Scale	71	0.76**
Positive and Negative Affect Scale – Positive Subscale	63	0.73*
Positive and Negative Affect Scale – Negative Subscale	63	-0.55**
Satisfaction With Life Scale	79	0.72**
Global Life Satisfaction Scale	77	0.55**
Scale of Psychological Well-being	63	0.73**
EQ-5D Thermometer	72	0.42**
Emotional Intelligence Scale	67	0.51**

$^{\alpha}$ Pearson's correlation coefficient

* => significant at 0.05 level

** => significant at 0.01 level

Correlations were moderately high between WEMWBS and the: Scale of Psychological Well-being; Satisfaction with Life Scale; Short Depression Happiness Scale; Positive and Negative Affect Scale – positive subscale; and the WHO-Five Well-being Index. These results were similar to those found between Affectometer 2 and these scales, which is as expected, given that Affectometer 2 was the starting point for research on the WEMWBS scale. These results indicate that WEMWBS covers both hedonic and eudaimonic aspects of mental well-being.

WEMWBS showed moderate to low correlations with the EQ-5D thermometer (a measure of overall physical and emotional health) and the Emotional Intelligence Scale (a measure of the ability to accurately assess one's own and others' emotions). This is expected because these two scales measure concepts that are separate from (but not unrelated to) positive mental health.

Internal consistency
Considers whether the scale describes a consistent underlying theme – in this case, it considers the extent to which WEMWBS's items are focused on assessing mental well-being. Scores range from 0 to 1 and are measured by Cronbach's alpha coefficient. The higher the co-efficient, the more highly correlated the items in the scale. A coefficient of 0.7-0.8 is ideal (Nunnally, 1978), and higher coefficients may suggest that some degree of item redundancy exists in the scale.

Cronbach's alpha coefficient = 0.89 (n = 348).

This high coefficient suggests that, while there is a good level of internal consistency, there may be scope to reduce even further the number of items in the scale (analyses are currently underway to explore the potential for a shortened scale, see section 10)

Test-retest reliability
Considers the stability of responses over a period of time. Test-retest reliability is determined by calculating the correlation between two sets of scores for the same group of people who repeat the test after a set period of time. For WEMWBS, the time period was one week.

Correlation$^{\alpha}$ = 0.83 after one week (n = 124)
$^{\alpha}$ Intra-class correlation coefficient

The test-retest reliability score was high for WEMWBS after one week. This suggests that the transient fluctuations that a person may experience from one day to the next are not reflected in the scores, and these scores remain robust over a short period of time.

Response Bias
Considers the extent to which an individual may tailor his or her responses in order to be perceived in a certain light, a phenomenon known as 'impression management'. And also the extent to which an individual remains unaware of

their true state of mental well-being known as 'self deception bias'. These two aspects of social desirability responding are measured using the Balanced Inventory of Desirable Responding (BIDR).

Correlations between the two subscales of the Balanced Inventory of Desirable Responding and WEMWBS, and between the two subscales and other mental health scales including Affectometer 2 are shown below:

Table 3: Correlation of WEMWBS to BIDR$^{\alpha}$

Scale	n	Impression Management	Self-Deception
WEMWBS	115	0.18*	0.35**
Affectometer 2	115	-0.25**	0.55**
WHO-Five Well-being Index	62	-0.39**	-0.20
Positive and Negative Affect Scales – Positive subscale	52	0.02	0.50**
Positive and Negative Affect Scales – Negative subscale	51	0.03	-0.16
Satisfaction with life scale	62	0.34**	0.40**
Global life satisfaction scale	62	0.26*	0.13

$^{\alpha}$ Pearson's correlation coefficient
* => significant at 0.05 level
** => significant at 0.01 level

WEMWBS showed a low correlation with both subscales of the Balanced Inventory of Desirable Responding. This contrasts with Affectometer 2, where self-deception bias was a major disadvantage of the scale. WEMWBS also performed better than three comparison mental health scales on impression management and better than two on self-deception.

These findings suggest that both impression management and self-deception response biases, whilst still an issue (as they are with all mental health scales), are acceptable for monitoring and evaluation purposes at the group/population level.

Focus groups

Face validity

Face validity assesses whether the items in the scale are suitable for the overall concept being measured. For WEMWBS this was tested in two mini focus groups with members of the general population in England and Scotland, selected on the basis of socioeconomic background, age and sex. Groups included mental health service users and non-users. Individuals were asked to complete WEMWBS and discuss their impressions of the scale. The aim of these investigations was to test what people thought WEMWBS was designed to measure and to determine its user-friendliness. Participants were asked to identify any items which they thought irrelevant or confusing. Results of these

focus group discussions suggested that WEMWBS was clear, user-friendly and unambiguous. Unlike the Affectometer 2, no suggestions were made to modify the scale or to clarify it in any way. Importantly, participants recognised that WEMWBS measured positive mental health rather than mental illness.

Scottish population samples (n =1749)

WEMWBS was included in the Autumn wave of the *Scottish Health Education Population Survey* (HEPS) 2006 (wave 12), which collected data from a random sample of the Scottish population aged 16 to 74 on a wide range of aspects of health and health related lifestyles (Gosling *et al.,* 2008), and also in the population survey '*Well? What do you think?*' (Well?) 2006, conducted on a random sample of the Scottish population aged 16 and above to collect data on public attitudes to mental health, mental well-being and mental health problems (Braunholtz *et al*., 2007).

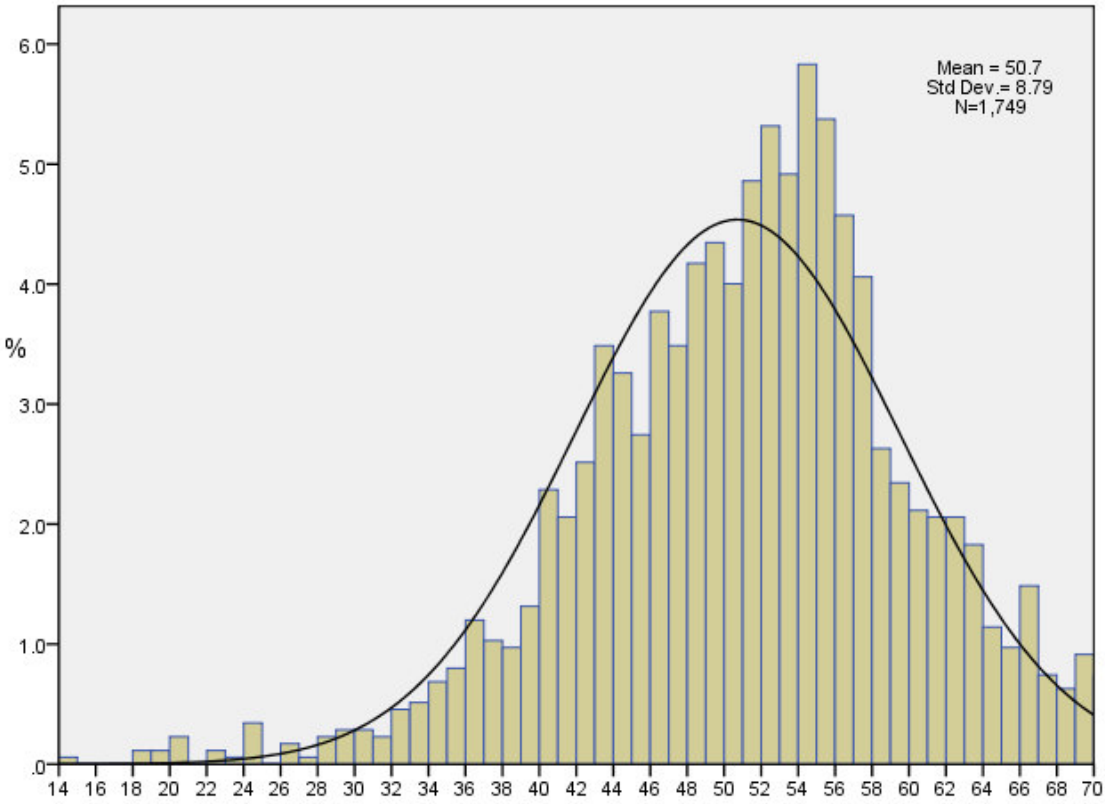
Analysis of combined data from these two population surveys (n = 2075 for the combined datasets, with complete WEMWBS scores for n = 1749 and complete GHQ 12 scores for n = 1239) have confirmed the findings of the student validation (Tennant *et al*., 2007):

- verification of a pre-hypothesised single underlying factor (n = 1749)
- Cronbach's alpha = 0.91 (n = 1749), again indicating that while there is a good level of internal consistency, there may be scope to reduce the number of items in the scale even further
- good performance against accepted criteria, discriminating population groups largely as expected and in a way consistent with other population surveys (see section 7 and Appendix iii)
- significant moderate negative correlation to the General Health Questionnaire 12 (GHQ 12) (see section 7).

5. Distribution of WEMWBS scores

In both the student and population samples, WEMWBS scores followed a roughly normal distribution with only a slight left-skew (Figure 2). WEMWBS can be used to calculate mean scores for different groups of people or for the same people at different time periods.[3] Mean scores can be compared using standard deviations and 95% confidence intervals.

Figure 2: Distribution of WEMWBS scores for the combined HEPS (wave 12) and Well? 2006 datasets (n = 1749)



Because WEMWBS scores show a roughly normal distribution, WEMWBS can be expected to capture the full spectrum of positive mental health without floor or ceiling effects and be suitable both for monitoring trends over time and evaluating the effect of mental health promoting programmes or interventions. However, although several studies are now in progress, it is important to note that, at the time of writing this manual, WEMWBS's sensitivity to change has not been demonstrated. As Affectometer 2 is sensitive to change there is no reason to think that WEMWBS will not be.

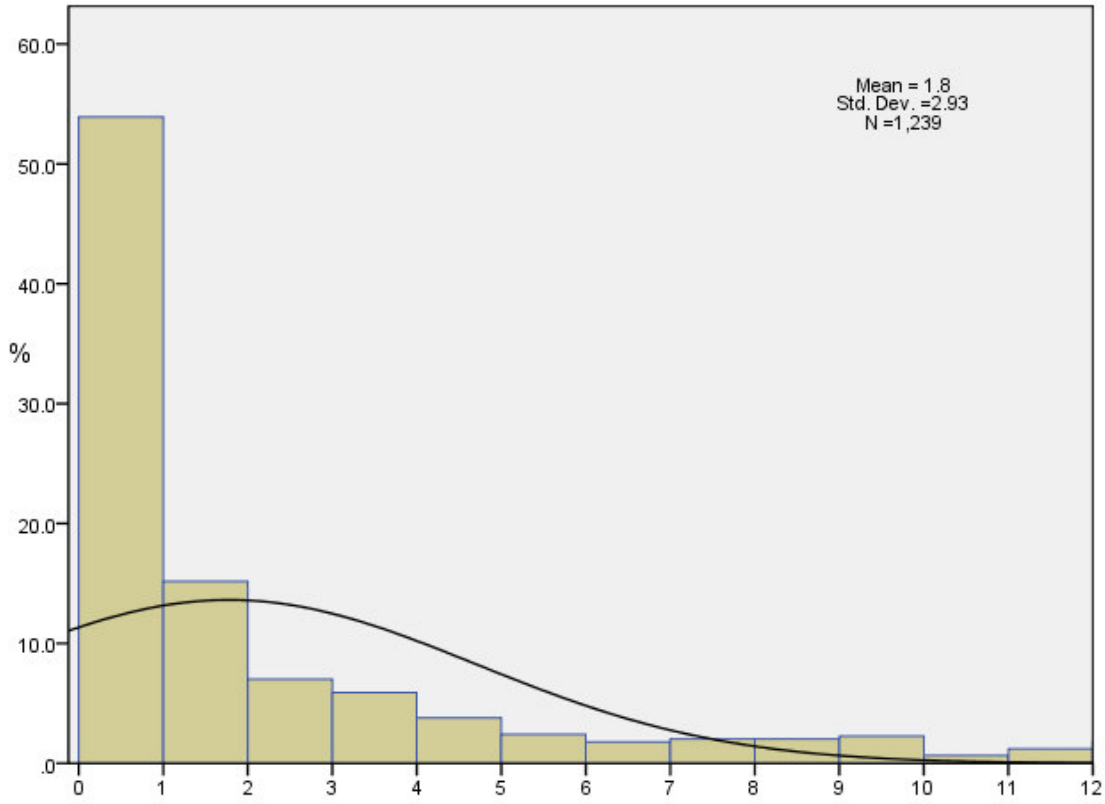
[3] Median scores should be used if data collected are not normally distributed, and mean scores if the data are. WEMWBS scores followed a roughly normal distribution with a slight left-skew. As the distribution is so close to normal it is considered appropriate to use mean scores, although some statisticians may decide that median scores should be used.

As well as not being designed to identify people who have or probably have a mental illness, WEMWBS does not a have a ‘cut off’ level to divide the population into those who have ‘good’ and those who have ‘poor’ mental well-being in the way that scores on other mental health measures, for example the GHQ 12 do (see section 6).

6. Comparison between WEMWBS scores and scores on the GHQ 12

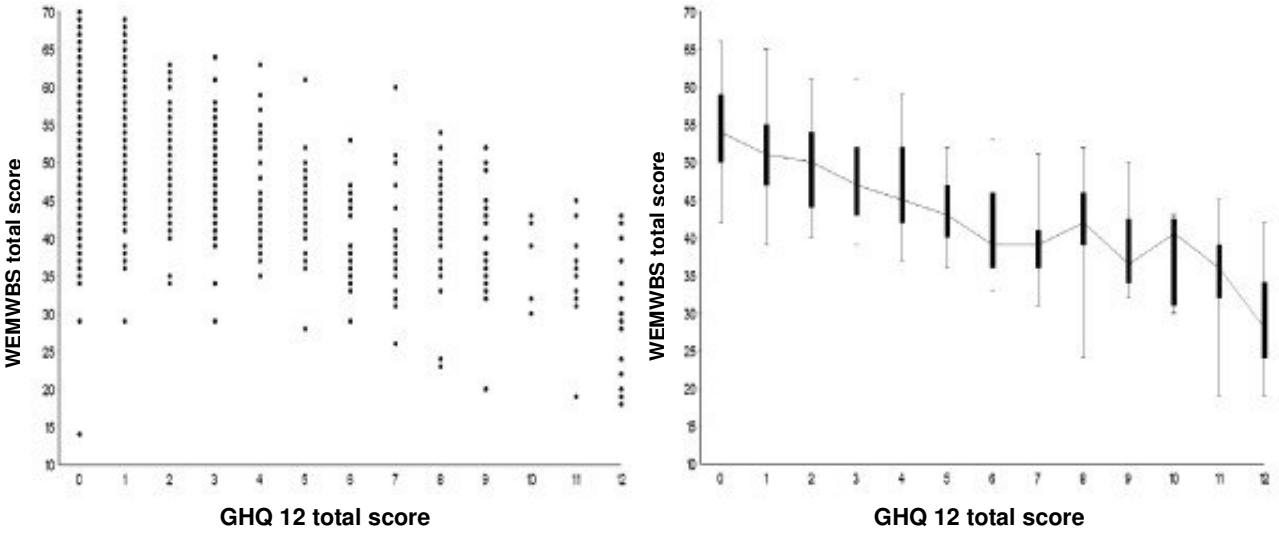
Both the HEPS and Well? surveys included the GHQ 12 measure alongside WEMWBS, allowing the two to be compared in the same group of people. As Figure 3 shows, GHQ 12 scores are heavily 'skewed' in population samples with the majority of people scoring 0 (no evidence of possible mental illness). This contrasts with the distribution of WEMWBS scores (Figure 2). Because of this distribution, GHQ 12 scores are more appropriately analysed in groups representing different levels of mental illness. A cut off score of 4+ is often used, with the 4+ group being more likely to have a diagnosable mental illness.

Figure 3: Distribution of GHQ 12 scores for the combined HEPS (wave 12) and Well? 2006 datasets (n = 1,239)



WEMWBS scores showed a significant moderate negative correlation with GHQ 12 scores in this population ($r = -0.53$, $p < 0.01$, Spearmans rank correlation), which persisted when a dichotomous scoring method, (with the four GHQ 12 response categories being scored 0, 0, 1, 1) was used ($p < 0.01$) (see Figure 4 scatterplot and box and whisker plot). The scatterplot (left) shows that respondents scoring the same on the GHQ 12 had a wide range of WEMWBS scores, so although lower WEMWBS scores tend to be associated with higher GHQ 12 scores (right), one is not simply the inverse of the other. The two scales are therefore not measuring the same thing.

Figure 4: WEMWBS score vs. GHQ 12 score, scatter plot and box and 90% confidence interval whisker plot: population sample



Tennant *et al. Health and Quality of Life Outcomes* 2007 **5**:63 doi:10.1186/1477-7525-5-63

7. Variation across demographic & social groups

Analysis of data from these two population surveys has also provided provisional population norms for WEMWBS across different socio-demographic groups. Table 4 below shows mean WEMWBS scores along with the lower and upper 95% confidence intervals and the number of responses on which these estimates are based (Appendix iii shows the same analysis but for median scores and also contains significance p values).

The provisional population mean score is 50.7 with 95% confidence interval 50.3 – 51.1 from the combined HEPS (wave 12) and Well? 2006 datasets.

Table 4: WEMWBS mean scores across demographic groups: population sample combined HEPS (Wave 12) and Well? 2006 datasets (n= 1,749)

Variable		n	Mean (95% CI)
Total		1749	50.7 (50.3 - 51.1)
Sex			
	Male	783	51.3 (50.6 - 51.9)
	Female	966	50.3 (49.7 - 50.8)
Age in years			
	16 – 24	176	51.7 (50.6 - 52.8)
	25 – 34	245	50.1 (49.1 - 51.1)
	35 – 44	353	49.7 (48.8 - 50.7)
	45 – 54	306	49.5 (48.4 - 50.5)
	55 – 64	334	51.4 (50.4 - 52.4)
	65 – 74	274	52.4 (51.3 - 53.4)
	75+	61	51.2 (48.9 - 53.4)
Tenure			
	Own outright	523	52.3 (51.5 - 53.0)
	Own with a mortgage	705	51.1 (50.5 - 51.7)
	Rent	519	48.6 (47.8 - 49.4)
Self-perceived health status			
	Very good	563	53.8 (53.1 - 54.5)
	Good	753	50.9 (50.4 - 51.9)
	Fair	319	47.6 (46.6 - 48.6)
	Poor	84	43.5 (41.3 - 45.6)
	Very poor	29	40.9 (37.1 - 44.6)
Employment Status			
	In work	968	51.4 (50.9 - 51.9)
	Student	82	51.8 (50.2 - 53.4)
	Retired	465	50.6 (49.8 - 51.4)
	Unemployed	154	48.4 (47.0 - 49.8)
	Other	79	46.1 (43.5 - 48.8)
Marital Status			
	Single	188	49.4 (48.2 - 50.7)
	Married/Living as couple	418	51.7 (50.9 - 52.5)
	Widowed/Divorced/Separated	155	47.8 (46.1 - 49.5)

Gross household income, pa		
<£5000	55	48.3 (46.0 - 50.6)
5000 – 14999	198	48.3 (46.9 - 49.7)
15000 – 29999	180	52.3 (51.1 - 53.5)
30000+	173	50.6 (49.3 - 51.8)
Terminal Education Age		
<16	228	50.7 (49.5 - 51.9)
16 – 18	355	49.2 (48.2 - 50.2)
19+	181	51.8 (50.7 - 52.9)
Chief Income Earner Social Grade		
A	38	52.7 (49.5 - 55.9)
B	84	50.68 (48.8 - 52.5)
C1	217	51.5 (50.5 - 52.6)
C2	193	51.0 (49.8 - 52.2)
D	101	49.5 (47.7 - 51.3)
E	124	46.8 (45.0 - 48.7)

95% CI = 95% confidence interval of the mean

In this large dataset small differences reach statistical significance (meaning that the differences are likely to reflect real differences in the population). Significant differences in mental well-being were found for each of the five categories of 'self perceived health status', ranging from very good to very poor. For tenure, those who rent were found to have significantly lower mental well-being scores from those who own outright and own with a mortgage. Those who were unemployed had significantly lower mental well-being scores than those who were in work or studying, although no significant differences were found between those who were retired compared to each of the other 4 employment categories. For marital status, those who were married or living as a couple had significantly higher mental well-being then those who were categorised as single or as widowed/divorced/separated. No real pattern was found for mental well-being with respect to gross household income per annum or terminal education age. There were no significant differences found either for chief income earner social grade, gender or age, although there appears to be a trend towards lower mental well-being for lower social grades and a U-shaped relationship for age.

These are the first results for WEMWBS and larger surveys are required before population norms are fully established. The availability of data on WEMWBS from, for example, the Scottish Health Survey (n = 6,000) from 2008 will help in this respect.

8. Using WEMWBS

WEMWBS is free to use but permission needs to be sought. Further information is included in Appendix i.

Data Collection
To date, WEMWBS has been administered in a self-completion format. This has been either via CASI (computer assisted self interviewing) whereby respondents are invited to enter their responses directly into the CAPI (computer assisted personal interview) machine (Well? survey and HEPS) or by the self-completion of paper formats of the scale (student samples and focus groups). WEMWBS can be assumed to be robust using either of these methods.

WEMWBS has not been tested in interview situations where an interviewer reads out the items to respondents and fills in their responses for them. We do not therefore know if WEMWBS is robust in these situations.

Scoring
Each of the 14 item responses in WEMWBS are scored from 1 (none of the time) to 5 (all of the time) and a total scale score is calculated by summing the 14 individual item scores (Table 5). The minimum score is 14 and the maximum is 70.

Table 5: Example: Scoring of WEMWBS - with responses highlighted in green

Statements	None of the time	Rarely	Some of the Time	Often	All of the time
I've been feeling optimistic about the future	1	2	3	4	5
I've been feeling useful	1	2	3	4	5
I've been feeling relaxed	1	2	3	4	5
I've been feeling interested in other people	1	2	3	4	5
I've had energy to spare	1	2	3	4	5
I've been dealing with problems well	1	2	3	4	5
I've been thinking clearly	1	2	3	4	5
I've been feeling good about myself	1	2	3	4	5
I've been feeling close to other people	1	2	3	4	5
I've been feeling confident	1	2	3	4	5
I've been able to make up my own mind about things	1	2	3	4	5
I've been feeling loved	1	2	3	4	5
I've been interested in new things	1	2	3	4	5
I've been feeling cheerful	1	2	3	4	5
Scores	0	0	4 x 3 = 12	4 x 4 = 16	6 x 5 = 30

Total Score = 0 + 0 + 12 + 16 +30 = 58

Presenting the results
WEMWBS results should be presented as a mean score for the population of interest with either a standard deviation or 95% confidence interval. The latter both provide a measure of variance of the scores in the population studied (either as a whole or for sub-groups within it). The range of scores within a sample can also be presented. Scores will vary between 14 and 70.

Interpreting the results
Table 4 (page 14) shows that the average population mean is around 51 and that this varies according to the population group studied. The mean score for the population under study can be compared with these provisional population norms to assess whether the level of mental well-being is above or below this level.

Differences between the scores of different groups or between the scores of the same group of people at two points in time, for example, before and after an intervention, need to be tested statistically using students t-Test or equivalent to assess how likely the differences are to have arisen by chance. At any given level of difference results are more likely to be significant if the groups being compared are large and less likely if the groups are small. A sample size calculation needs to be carried out to assess how big a group should be to show statistical significance of a specific difference. Table 6 gives examples of this for different sample sizes using WEMWBS data from the HEPS and Well? surveys combined.

Table 6: Examples of sample size required

Population size	Difference in WEMWBS scores between two groups			
	± 1 points	± 2 points	± 3 points	± 5 points
10,000	1082	294	133	48
50,000	1184	301	134	48
100,000	1199	302	135	49

Sample size (per group) based on difference in mean scores of two groups using a power of 0.8, a significance level of 0.05 and population sample combined HEPS (Wave 12) and Well? 2006 datasets (n = 1,749).

If groups within the sample are to be compared, then the sample size calculation needs to be based on these groups, for example, men separately from women, and not on the total sample size ie the men plus women.

Dealing with missing data
For the WEMWBS validation, HEPS and Well? responders were deleted if they were not full-responders (ie they did not answer all items of WEMWBS). This harsh method was appropriate as the vast majority of responders were full-responders and thus loss of sample size was minimal. However, it may be too harsh an approach to adopt in other surveys.

Views differ on how to deal with missing data and none of the possible methods have been assessed for WEMWBS. The problem of missing data in multi-item

scales is curiously under-discussed in the methodological literature. Some researchers use estimation to 'fill in' missing values, thus retaining their original sample size. The following are noted as alternative methods, to deleting respondents who are not full-responders, that have been suggested in literature:

- calculating the mean value of responses to items that a respondent has answered, and then using that mean score as the score for those questions which that respondent did not answer.
- using the midpoint of the range of possible responses
- using the mean response for the particular item from all respondents

However, using estimations to fill in missing values should only be done in situations where at least a certain proportion of items are answered. If less than this proportion has been answered the respondent's score should be set to missing. Researchers do not agree on what the proportion should be. For WEMWBS it can be anticipated that estimations for more than three missing items is unlikely to be robust. In such cases, the WEMWBS score should therefore not be calculated and should be set as missing. It will also be important to check the 'randomness' of the missing data to ensure that certain items are not being systematically missed. Overall, however, when dealing with missing WEMWBS data it is important to note that the effect of using estimations for WEMWBS scores has not been tested.

Using WEMWBS in individuals

WEMWBS provides robust results for populations and groups. It has not yet been validated for monitoring mental well-being in individuals.

9. Current usage of WEMWBS

WEMWBS is currently being used in numerous surveys and intervention studies, for example:

Surveys

- Scottish Health Survey from 2008
- Scottish Prison Service 2007 Annual Prison Survey
- British Social Attitudes Survey 2007
- HEPS Spring wave (wave 13) 2007 (8-item version in Autumn wave (wave 14) 2007)
- NHS Grampian population surveys
- A large population survey in Iceland
- National Childhood Development Study 2008 sweep
- Under consideration for the Scottish Household Survey 2009
- Under consideration for the UK Household Longitudinal study 2009
- Under consideration for the Health Survey for England

Local evaluations

- Evaluations of local Arts on Prescription Services
- Assessments of social prescribing projects
- Evaluating the impact of parenting programmes on parents' mental well-being in the Parenting Interventions Evaluation of Pathfinders projects
- Evaluation of Lottery funded projects in England

Monitoring interventions

- Monitoring mental well-being among patients attending psychiatric day hospital
- Occupation therapy interventions at a day hospital

WEMWBS is also being used as a **national indicator** in the:

- Scottish Government's Strategic Outcome Indicators 2008 for the Scottish Government's performance framework to monitor the spending review (Scottish Health Survey data from 2008) (Scottish Government, 2007)
- Scottish mental health and well-being indicator set (Scottish Health Survey data from 2008)

And is being considered by the Department for Environment, Food and Rural Affairs (DEFRS) for one of its indicators of well-being, specifically positive mental health, in its sustainable development indicators (Defra 2007).

10. Further validation research on WEMWBS

Although the basic establishment of WEMWBS is now completed, further research is being undertake in a number of areas as indicated below.

The scaling properties of WEMWBS
Data from the HEPS (wave 12) 2006 and Well? 2006 population surveys are being used to establish the extent to which WEMWBS fits the Rasch model. This is a statistical procedure used to determine how the intervals in an ordinal scale relate to one another. This analysis permits an answer to the following question: "Is a score of 60 twice as good as a score of 30?". A good fit to the Rasch model indicates that the scale has good scaling properties. This means that a mean score of, for example, 44 can be taken to be twice as good as a mean score of 22. Rasch analysis can also be used to determine potential item redundancy in a scale and to assess whether a reduction in the number of scale items may be appropriate.

Initial indications suggest that it may be possible to develop a shortened (seven or eight item) version of WEMWBS which has more robust scaling properties than the full scale.

WEMWBS's sensitivity to change
Several studies are ongoing to assess the sensitivity to change of WEMWBS. For example, WEMWBS is being used to audit the mental well-being of people attending a psychiatric day hospital. Changes on WEMWBS will be compared with patient and clinical assessments of mental health at the beginning and end of each patient's admission. If WEMWBS scores change significantly in the direction indicated by clinical assessment, sensitivity to change will be confirmed.

WEMWBS is also being used to assess mental well-being in parents attending a range of parenting programmes in the UK before they embark on the programme and at the end of the programme. These programmes have previously been shown to have a positive impact on parents' mental health and demonstration of changes in WEMWBS scores will confirm that the measure is sensitive to change.

Validation of WEMWBS with secondary school children aged 13 to 15 years
Research began in March 2008 to establish whether WEMWBS can be used to assess the overall mental well-being of children of secondary school age (13 to 15 years of age). This is the 'Warwick-Edinburgh Mental Well-being Scale Acceptability and Validation in English and Scottish Secondary School Students project' (The WAVES Project) being undertaken by Warwick and Edinburgh Universities. This will report October 2011.

Other research required
Other research still required includes assessing the extent to which it is appropriate to use WEMWBS to assess mental well-being among different ethnic

minority populations in the UK, and other cross-cultural validation for use of WEMWBS in countries other than the UK.

Further ahead

As understanding of mental well-being develops over the next decade, it is likely that measurement scales will also need to evolve. Whilst WEMWBS fulfils criteria for monitoring mental well-being at present and represents a very significant step forward in terms of other currently available measures, it is likely that it will need to undergo further development in the future.

This manual will be updated as results of the continuing validation of WEMWBS are known.

Bibliography

Braunholtz S, Davidson S, Myant K and O'Connor R (2007). *Well? What do you think?* (2006): The third national Scottish survey of public attitudes to mental health, mental wellbeing and mental health problems. Scottish Executive: Edinburgh. To access the dataset contact Angela Hallam Angela.Hallam@scotland.gsi.gov.uk

Defra (2007). *Sustainable development Indicators in your pocket 2007* – An update of the UK Government Strategy indicators. Department for Environment, Food and Rural Affairs: London.

Kammann R, and Flett R (1983). Affectometer 2: A scale to measure current level of general happiness. *Australian Journal of Psychology,* 35 (2), 259-265.

Gosling R, Bassett C, Gilby N, Angle H, and Catto S (2008). *Health Education Population Survey: Update from 2006 survey*. NHS Health Scotland: Glasgow. [database on the Internet]. Colchester: UK Data Archive; 2007. Available from: http://www.data-archive.sc.uk

Nunnally JC (1978). *Psychometric Theory*. Second Edition. McGraw-Hill: London.

Ryan RM, and Deci EL (2001). On happiness and human potential: a review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52*,* 141-166.

Scottish Government (2007). *Scottish budget spending review 2007*. The Scottish Government: Edinburgh.

Tennant R, Fishwick F, Platt S, Joseph S., and Stewart-Brown S (2006). *Monitoring Positive Mental Health in Scotland: Validating the Affectometer 2 Scale and Developing the Warwick-Edinburgh Mental Well-being Scale for the UK*. NHS Health Scotland: Glasgow.

Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, Parkinson J, Secker S, and Stewart-Brown S (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health & Quality of Life Outcomes,* 5 (63), doi:10.1186/1477-7525-5-63.

Tennant R, Joseph S, and Stewart-Brown S (2007). The Affectometer 2: a measure of positive mental health in UK populations. *Quality of Life Research,* 16 (4), 687-695.

Appendix i

Conditions of Using WEMWBS

We welcome the use of WEMWBS. It is free to use but is copyrighted to NHS Health Scotland and the Universities of Warwick and Edinburgh. Permission is required for use. Dr Kulsum Janmohamed K.Janmohamed@Warwick.ac.uk, working with Professor Sarah Stewart-Brown at the University of Warwick, is maintaining a register of use and is the person to contact when seeking such permission.

When you seek permission for use you should indicate how you are planning to use WEMWBS. We ask that after use you feed back to Dr Janmohamed on how WEMWBS has performed. Dr Janmohamed is also the person to contact should you have more questions regarding the scale and its use.

If the scale is reproduced, it must include the copyright statement which appears below it and no changes to its wording, response categories or layout must be made.

Any report regarding use of WEMWBS should include the following text:

"The Warwick-Edinburgh Mental Well-being Scale was funded by the Scottish Government National Programme for Improving Mental Health and Well-being, commissioned by NHS Health Scotland, developed by the University of Warwick and the University of Edinburgh, and is jointly owned by NHS Health Scotland, the University of Warwick and the University of Edinburgh."

The Warwick-Edinburgh Mental Well-being Scale (WEMWBS)

Below are some statements about feelings and thoughts.

Please tick the box that best describes your experience of each over the last 2 weeks

STATEMENTS	None of the time	Rarely	Some of the time	Often	All of the time
I've been feeling optimistic about the future	1	2	3	4	5
I've been feeling useful	1	2	3	4	5
I've been feeling relaxed	1	2	3	4	5
I've been feeling interested in other people	1	2	3	4	5
I've had energy to spare	1	2	3	4	5
I've been dealing with problems well	1	2	3	4	5
I've been thinking clearly	1	2	3	4	5
I've been feeling good about myself	1	2	3	4	5
I've been feeling close to other people	1	2	3	4	5
I've been feeling confident	1	2	3	4	5
I've been able to make up my own mind about things	1	2	3	4	5
I've been feeling loved	1	2	3	4	5
I've been interested in new things	1	2	3	4	5
I've been feeling cheerful	1	2	3	4	5

"Warwick Edinburgh Mental Well-Being Scale (WEMWBS)
© NHS Health Scotland, University of Warwick and University of Edinburgh, 2006, all rights reserved."

Appendix ii

Description of scales used to assess the construct validity of WEMWBS

Scales of affect/feelings	
WHO-Five Well-being Index (WHO-5)	Five item scale of statements covering key mental affect states (e.g. I felt cheerful, calm, vigorous and interested) with 5 response categories. All items positively worded.
Short Depression Happiness Scale (SDHS)	Six item scale with 4 response categories focused on affect with balanced positive and negative items (e.g. I felt happy, I felt cheerless)
Positive and Negative Affect Scale (PANAS): Positive subscale (PANAS P) Negative subscale (PANAS N)	Twenty item scale with 5 response categories comprising a list of positive and negative adjectives covering a wider variety of feelings than is usual in mental health scales (e.g. ashamed, attentive, proud, guilty, and excited).

Scales of subjective well-being	
Global Life Satisfaction Scale (GLSS)	Single item scale with 4 point response category. 'On the whole are you satisfied with your life?' Most commonly used measure of subjective well-being.
Satisfaction with Life Scale (SWLS)	Five item scale with 7 response categories. Items cover positive statements e.g. 'in most ways my life is close to ideal'. The prototype measure of well-being.

Psychological functioning	
Scales of Psychological Wellbeing (SPW)	Fifty four item scale with 6 response categories assessing psychological functioning with subscales measuring autonomy, self acceptance, environmental mastery, purpose in life, personal growth and positive relations with others.

Emotional Intelligence	
Emotional Intelligence Scale (EIS)	Thirty three item scale with 5 response categories. Consists of statements covering appraisal, expression, and regulation of emotion in self and others, and the utilisation of emotions in

	problem solving.

Psychiatric Morbidity	
General Health Questionnaire 12 (GHQ 12)	Twelve-item scale with 4 response categories. A well-established screening instrument designed to detect possible psychiatric morbidity in the general population. Respondents are asked to respond to questions relating to their recent experience of anxiety, self-confidence ability to concentrate, decision-making capacity, enjoyment of day-today activities, sleep disturbance and stress etc.

General Health	
EQ-5D thermometer	A measure of health in general where respondents rate their overall health (physical and mental) on a 0-100 scale. Responses to this scale tend to reflect physical more than mental health.

Response Bias	
Balanced Inventory of Desirable Responding (BIDR)	Forty-item scale, split into two sub-scales. The first sub-scale measures self-deception (SD) (the tendency to exaggerate certain responses or behaviours) and the second sub-scale measures impression management (IM) (the tendency to over-report desirable behaviours and under-report undesirable behaviours).

Appendix iii

WEMWBS median scores across demographic groups: population sample Combined HEPS (Wave 12) and Well? 2006 Datasets

Variable		N	Median (95% CI)	p
Total		1749	51 (51-52)	
Sex				
	Male	783	52 (51-52)	<0.05
	Female	966	51 (50-52)	
Age in years				
	16 – 24	176	53 (52-53)	<0.01[KW]
	25 – 34	245	51 (50-53)	
	35 – 44	353	51 (49-52)	
	45 – 54	306	50 (49-51)	
	55 – 64	334	52 (51-53)	
	65 – 74	274	52 (51-54)	
	75+	61	51 (49-54)	
Tenure				
	Own outright	523	52 (52-53)	<0.01[KW]
	Own with a mortgage	705	52 (51-52)	
	Rent	519	50 (49-51)	
Self-perceived health status				
	Very good	563	54 (54-55)	<0.01[J]
	Good	753	51 (51-52)	
	Fair	319	47 (46-49)	
	Poor	84	44 (40-46)	
	Very poor	29	41 (36-47)	
Employment Status ^				
	In work	968	52 (51-52)	<0.01[KW]
	Student	82	52 (50-54)	
	Retired	465	51 (50-52)	
	Unemployed	154	49 (47-51)	
	Other	79	46 (43-50)	
Marital Status *				
	Single	188	51 (49-53)	<0.01[KW]
	Married/Living as couple	418	52 (51-53)	
	Widowed/Divorced/Separated	155	49 (46-51)	
Gross household income, pa *				
	<£5000	55	48 (44-53)	<0.01[J]
	5000 – 14999	198	49 (47-51)	
	15000 – 29999	180	53 (51-54)	
	30000+	173	51 (49-53)	
Terminal Education Age *				
	<16	228	52 (50-53)	<0.05[KW]
	16 – 18	355	50 (49-51)	
	19+	181	53 (51-54)	
Chief Income Earner Social Grade *				
	A	38	55 (51-57)	<0.01[J]
	B	84	50 (48-53)	
	C1	217	51 (50-53)	
	C2	193	53 (51-54)	

D	101	50 (47-52)
E	124	47 (44-51)

* Tests conducted on a reduced set of individuals. Variable only recorded in the HEPS survey.
95% CI = 95% confidence interval of the median
^{KW} = p-value generated from a Kruskal-Wallis test.
^{J} = p-value generated from a Jonckheere's tests for ordered alternatives.
^ = test conducted excluding the Other category